Infiniband und Virtualisierung

Ulrich Hamm uhamm@cisco.com
Data Center Team

© 2005 Cisco Systems, Inc. All rights reserved.

Infiniband Primer

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

What is InfiniBand?

- InfiniBand is a high speed – low latency technology used to interconnect servers, storage and networks within the datacenter
- Standards Based – InfiniBand Trade Association

  http://www.infinibandta.org
- Scalable Interconnect:

  1X = 2.5Gb/s

  4X = 10Gb/s

  12X = 30Gb/s

Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners
3

InfiniBand Physics

- Copper and Fibre interfaces are specified
- Copper

  Up to 20m for 4x connections

  Up to 10m for 12x connections
- Optical

  Initial availability via dongle solution

  Up to 300m with current silicon (probably 150m first)

  Long Haul possible, but not with current silicon

Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners
4

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand Physics

- **Link is bonded 2.5Gbps (1x) links**
  - **Fiber is a ribbon cable**
  - **Copper is a multi-conductor cable**
- **Each Link is 8b/10b encoded**
  - **4x Link is 4 2.5Gbps Physical Connections**
  - **Each connection is 2Gbps data**
  - **SAR provides a single 8Gbps data connection (4x)**
    - **24 Gbps (12x)**

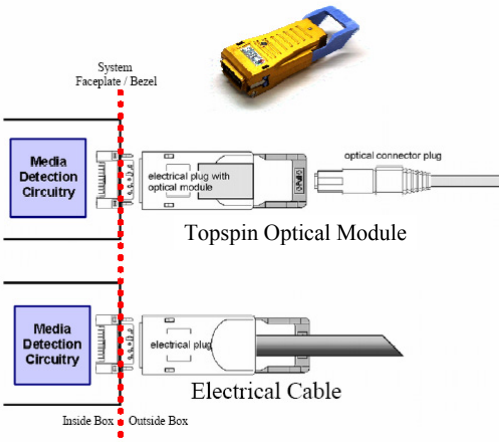
© 2005 Cisco Systems, Inc. All rights reserved.

Pluggable Optics Module

Transforms Powered Copper Ports to Optical Ports

Coverts a copper port to an optical port *on a port by port basis*

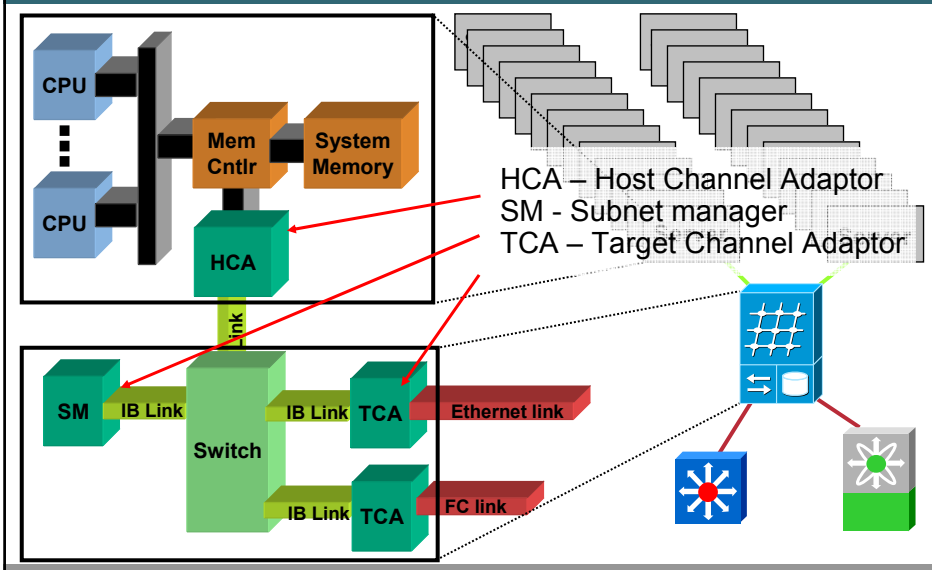
Extends port to port reach to 150m - 300m with fibre ribbon cables



© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand Nomenclature



© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand Switch Hardware

- **Hardware switch devices is a cut-through memory switch**
- **Full-duplex, non-blocking tag forwarding switch**
- **Tags are system Local ID, provided to all network endpoints by the Master Subnet Manager on system startup**

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand Host Channel Adapter

- Network interface for IB attached Servers
- Provides hardware Virtual/Physical memory mapping, Direct Memory Access (DMA), and memory protection
- Provides RDMA (Remote DMA) data transfer engine and reliable packet forwarding capabilities

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand Gateway

- Technically a Target Channel Adapter
- Similar to an HCA attached to an embedded device
- Usually doesn't require virtual memory manipulation and mapping
- Simplified HCA on a specialized device

  Examples, Ethernet to InfiniBand or Fibre Channel to InfiniBand packet fowarding engines

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand System Overview

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand System Architecture

- **Connection Oriented Architecture**
  - **Central connection routing management (SM)**
  - **All communications based on send/receive queue pairs**
- **Two primary connection types**
  - **Reliable Connection**
  - **Unreliable Datagram**
- **Unused connection types**
  - **Unreliable Connection**
  - **Reliable Datagram**
  - **Raw Datagram**

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand Connections

- Reliable Connection
  - Host Channel Adapter based guaranteed delivery
  - Uses HCA onboard memory (or system memory with PCI-E) for packet buffering
  - Primarily used for RDMA communications
  - Can use end-to-end flow control based on credits related to available receive buffers
- Unreliable Datagram
  - Best effort forwarding
  - Used for IP over IB communications

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand Subnet Manager

- IB Fabric is called an InfiniBand Subnet
  - All devices under the control of a single Master Subnet Manger (SM)
  - May have multiple slaves with replicated SM database state
- At system startup, all devices register with the SM
  - Central Routing function
    - Shortest Path First Routing
    - Equal Paths Loadbalanced with static round robin distribution
  - Connection endpoint lookup

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Clusters 2.0 Subnet Manager: Fabric Sweep Performance

Number of Hosts	Time
32	< 1 sec
64	< 1 sec
128	2 sec
256	4 sec
512	22 sec
1,024*	35-40 sec
2,048*	1-1:30 min**
4,096*	5-7 min**

** Requires HPC Subnet Manager for this performance*
*** Estimated based on simulation*

- Assumes InfiniSwitch-III based two tier topology
- Embedded SM can handle up to 1,024 nodes

© 2005 Cisco Systems, Inc. All rights reserved.

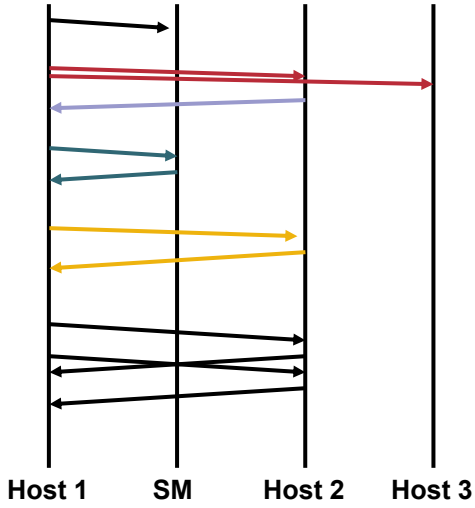
IB Addressing

- **3 addresses: GUID, GID, LID**
- **GUID**
  - **Global Unique ID 64 bits in length**
  - **Used to uniquely identify a port or port group**
  - **HCA and each port has a GUID**
  - **(e.g 00:05:ad:00:00:01:02:03)**
- **GID**
  - **GUID plus Subnet prefix**
  - **Used for host lookup on a subnet**
  - **Used for inter-subnet IB routing (future)**
  - **(e.g. fe:80:00:00:00:00:00:00:05:ad:00:00:01:02:03)**

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Address Resolution



Join Multicast Group

Send ARP on “Broadcast addr”

Receive remote GID via ARP

Ask SM for GID->LID mapping

Ask Host for Service info (QP)

Communicate

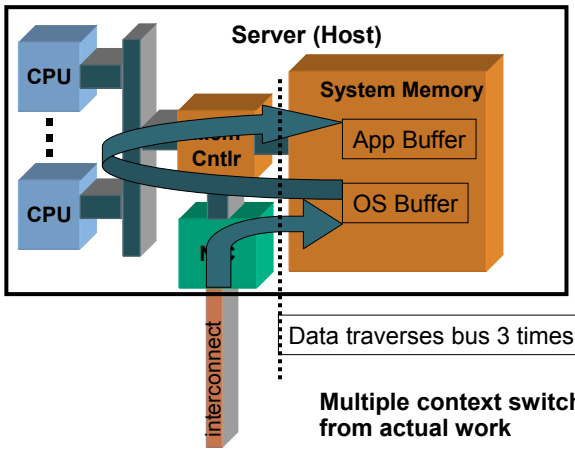
© 2005 Cisco Systems, Inc. All rights reserved.

RDMA and Upper Layer Protocols

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Current NIC Architecture



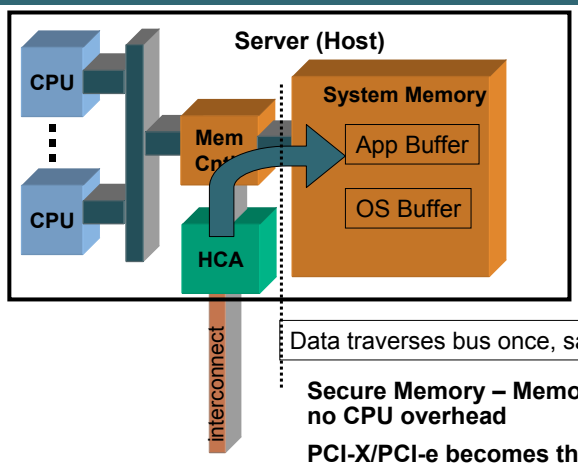
Multiple context switches robs CPU cycles from actual work

Memory bandwidth and per packet interrupts limit max throughput

OS manages end-to-end communications path

© 2005 Cisco Systems, Inc. All rights reserved.

With RDMA and OS Bypass



Secure Memory – Memory transfers with no CPU overhead

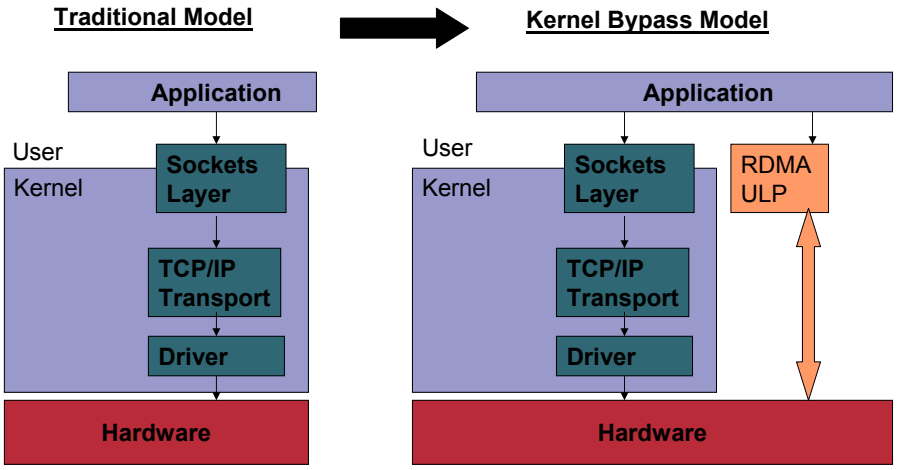
PCI-X/PCI-e becomes the bottleneck for network data transmission

HCA manages remote data transmission

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Kernel Bypass



© 2005 Cisco Systems, Inc. All rights reserved.

Upper Layer Protocols

- **Variety of software protocols to handle high speed communication over RDMA**
- **Protocols include**
  - **IP-over-InfiniBand – IETF** http://www.ietf.org/internet-drafts/draft-ietf-ipoib-ip-over-infiniband-09.txt
  - **SDP – InfiniBand Trade Association** http://infinibandta.org
  - **SRP – ANSI T10** http://www.t10.org/ftp/t10/drafts/srp/srp-r16a.pdf
  - **DAPL – DAT Collaborative** http://www.datcollaborative.org
  - **MPI – MPI Forum** http://www.mpi-forum.org

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

IP over InfiniBand

IPoIB

- IETF draft specification
- Leverages InfiniBand Multicast for broadcast requirements (ARP)
- Supports TCP, UDP, IP Multicast

© 2005 Cisco Systems, Inc. All rights reserved.

Sockets Direct Protocol

SDP

- STREAM Sockets over InfiniBand Reliable Connections
- TCP offload function for IB attached devices
- Can be used by TCP application without re-building the application
- Asynchronous I/O model also available with true RDMA forwarding – requires application re-write

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

SCSI RDMA Protocol

SRP

- SCSI Semantics over RDMA fabric
- Not IB specific
- Host drivers tie into standard SCSI/Disk interfaces in kernel/OS
- Can be used for end-to-end IB storage (implemented today!)

© 2005 Cisco Systems, Inc. All rights reserved.

Direct Access Provider Library

uDAPL

- Two variants: User DAPL (uDAPL)/Kernel DAPL (kDAPL)
- RDMA semantics API
- Provides low level interface for application direct or kernel direct RDMA functions (memory pinning, key exchange, etc.)

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Message Passing Interface

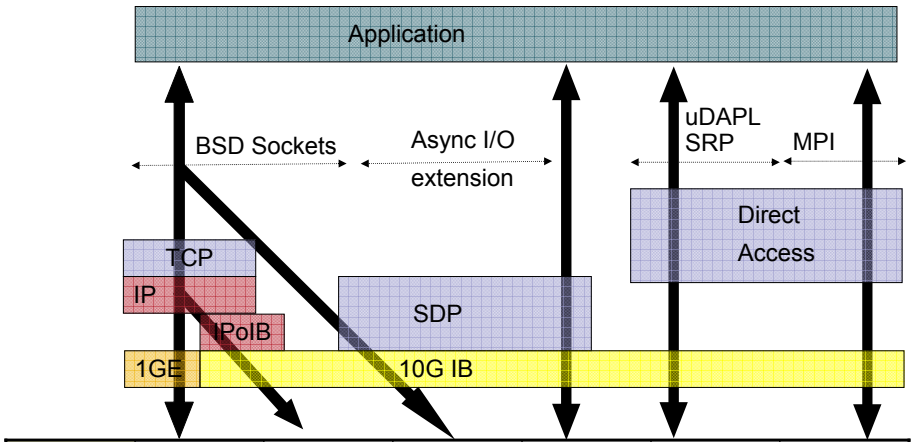
MPI

- **MPI is the defacto standard API for parallel computing applications**
- **RDMA capabilities added via a set of patches to the base MPI code (MPICH, one of many available MPI libraries), initially developed at Ohio State University**

  http://nowlab.cis.ohio-state.edu/projects/mpi-iba/

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand Performance
Measured Results



Throughput	1 Gb/s	4.1Gb/s	4.5 Gb/s	7.9Gb/s	8Gb/s	8Gb/s
	0.8Gb/s	1.2Gb/s	3.2Gb/s	6.2Gb/s	6.4Gb/s	
Latency	40-60 usec	30 usec	18 usec	18 usec	8 usec	3.5 usec

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

IB Glossary

- IB – InfiniBand Architecture (not InfinityBand)
- HCA – Host Channel Adapter (NIC)
- RDMA – Remote Direct Memory Access
- SM – Subnet Manager (management process)
- SRP – SCSI RDMA Protocol
- SDP – Sockets Direct Protocol
- TCA – Target channel Adapter (gateway)

© 2005 Cisco Systems, Inc. All rights reserved.

Why InfiniBand for HPC?

1. InfiniBand Technology Provides Industry Best Price / Performance and network efficiency
2. Mission Critical Management and Reliability Features for HPC Clusters
3. Industry standard with strong open source community and Linux adoption
4. Customer and application proof points: InfiniBand is a hot technology right now

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Cisco HPC Product Overview

Cisco SFS 7000, 7008, 7012, and 7024

© 2005 Cisco Systems, Inc. All rights reserved.

What Makes The Server Fabric Switch Different?

High Performance Server-to-Server Interconnect

Virtualization (I/O, Storage, and CPU)

Policy-Based Dynamic Resource Mapping

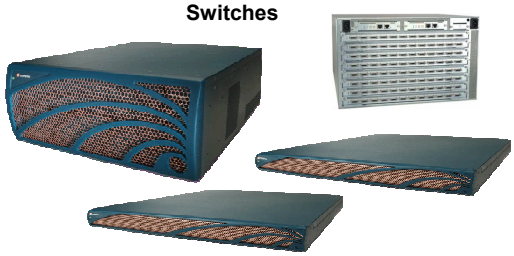
Performance and Control

© 2005 Cisco Systems, Inc. All rights reserved.

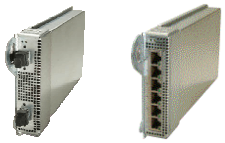
Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

SFS Building Blocks

Switches



Gateway Modules
- InfiniBand to Ethernet
- InfiniBand to Fibre Channel



Host Channel Adapter (HCA)
With upper layer protocols

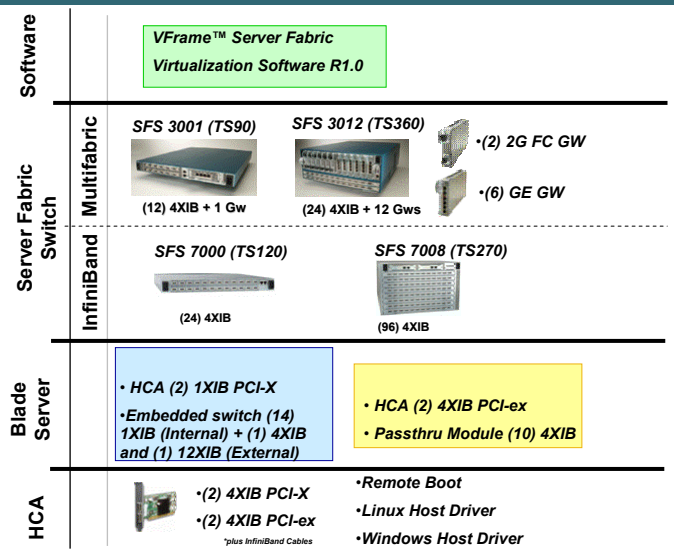
- SRP
- SDP
- uDAPL
- MPI
- IPoIB

Integrated System and Fabric management

Linux and Windows driver support

© 2005 Cisco Systems, Inc. All rights reserved.

The Cisco SFS Product Line



© 2005 Cisco Systems, Inc. All rights reserved.

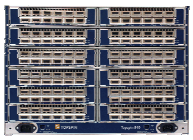
Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Terms and Components

Topspin IB Components

InfiniBand Switches

SFS 7000 (24-ports)
Generally used as an edge switch.



SFS 7008 (up to 96-ports)
SFS 7012 (up to 144-ports)
SFS 7024 (up to 288-ports)

Generally used as core switches in large fabrics, or as an edge switch in single switch fabrics.

Topspin InfiniBand Cable
Plugs into a switch either as a host or uplink cable.

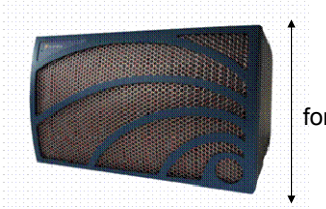
Topspin 2-port Low-Profile HCA
PCI-X and PCIe flavors available

Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners

SFS 7008 – 96-Port Core Switch

***Distributed* Building Block for Highly Manageable HPC Clusters**

- Rapid Service™ Architecture
  - Passive midplane
  - 2 min. service vs. 2 man days
- No single point of failure
  - Hot-swap power, fan, mgt
- Full security
  - SSH/SSL/SNMP v3, RADIUS, SCP
  - Multi-tiered Security Partitions
- Architected for Rolling upgrades for hitless service windows



6U
form factor

96 4X or 32 12X InfiniBand ports

8 Line Card Slots (8 x 12 = 96)

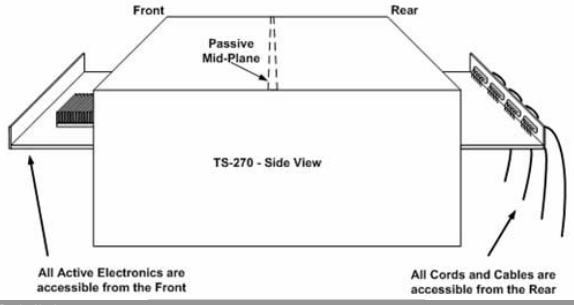
Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

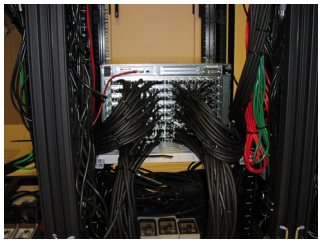
Mission Critical RAS Functionality

Cisco InfiniBand Rapid Service Architecture

- **InfiniBand switches that enable enough enterprise-class reliability, but still priced for high-port count HPC**
- **Full set of software and hardware diagnostics with active alerts**
- **Repair at FRU-level, not switch-level with minimal fabric disruption**


Front
Rear
Passive Mid-Plane
TS-270 - Side View
All Active Electronics are accessible from the Front
All Cords and Cables are accessible from the Rear

SARA



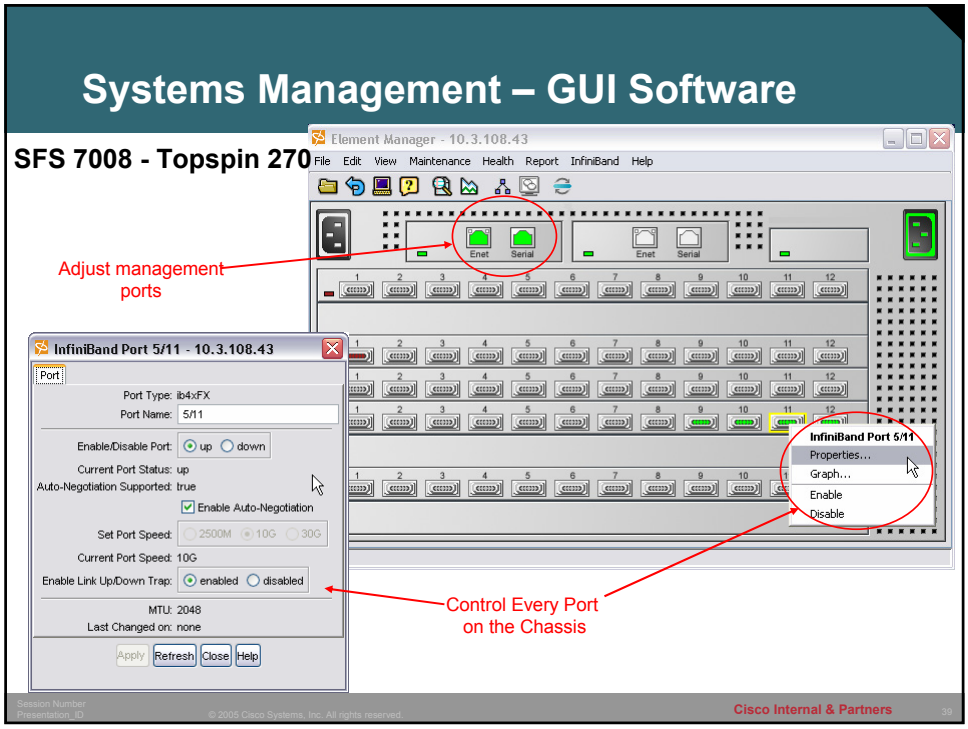
© 2005 Cisco Systems, Inc. All rights reserved.

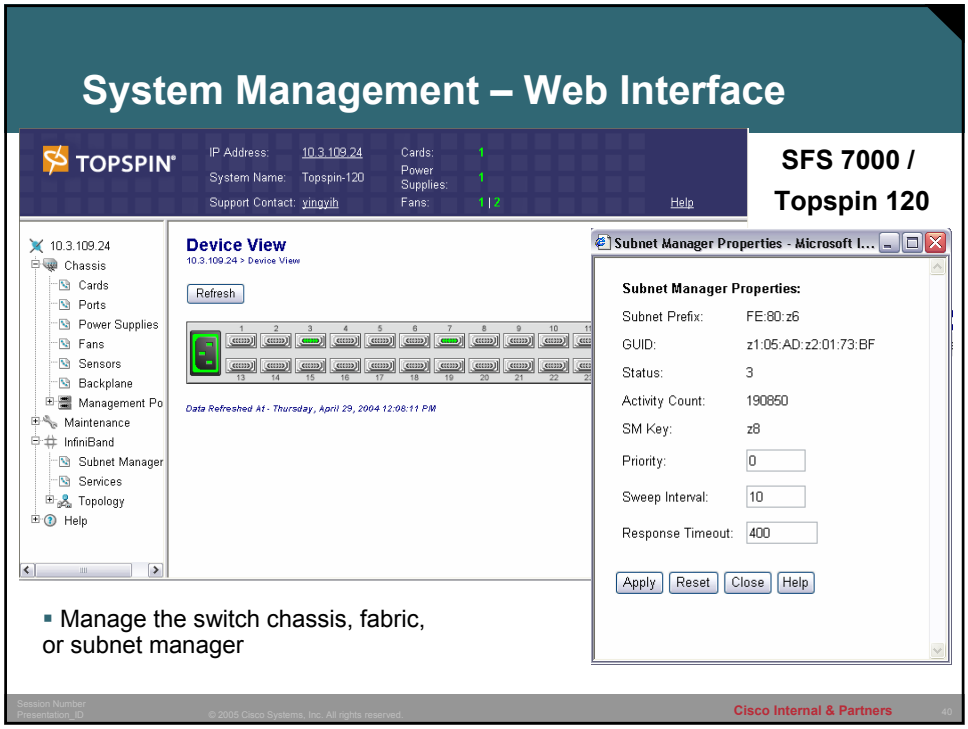
Cisco SFS Large Switch Family Overview

	SFS 7008	SFS 7012	SFS 7024
Chassis Type	6U Modular	7U Modular	14U Modular
Max 4X ports	96 ports	144 ports	288 ports
Max 12X ports	32 ports	48 ports	96 ports
Port Module Options	8 Horizontal Slots 12 by 4X LIMs 4 by 12X LIMs	12 Side by Side Slots 12 by 4X LIMs 4 by 12X LIMs	24 Side by Side Slots 12 by 4X LIMs 4 by 12X LIMs
High Availability	Redundant Power/Cooling Redundant Management Non-disruptive Upgrades Hot Swappable FRUs	Redundant Power/Cooling Redundant Management Non-disruptive Upgrades Hot Swappable FRUs	Redundant Power/Cooling Redundant Management Non-disruptive Upgrades Hot Swappable FRUs
Fabric Manager	Embedded or External	Embedded or External	Embedded or External
Best Use	▪ 64-96 node clusters ▪ Core switch for 288 to 4,000 node clusters	▪ 97-144 node clusters	▪ 145-288 node clusters ▪ Core switch for 1,536+ node clusters

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr





Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

I/O Virtualization

© 2005 Cisco Systems, Inc. All rights reserved.

The Evolution of I/O Virtualization

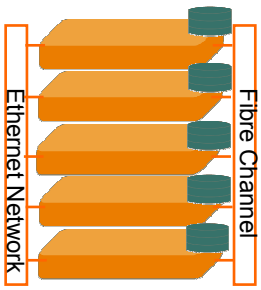
SMP

Symmetric Multi-Processor

Pro: Single managed entity, fast backplane

Con: Expensive, Proprietary server + backplane

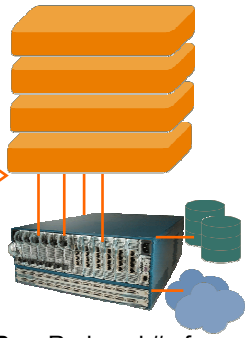
Dis-aggregation



Pro: Standard servers, inexpensive

Con: Lots of managed components, low-performing interconnect

Virtualization

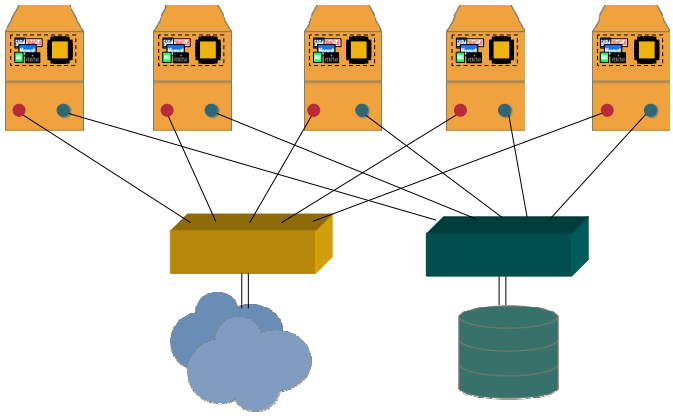


Pro: Reduced # of managed components, virtual I/O, *fast* standards backplane

© 2005 Cisco Systems, Inc. All rights reserved.

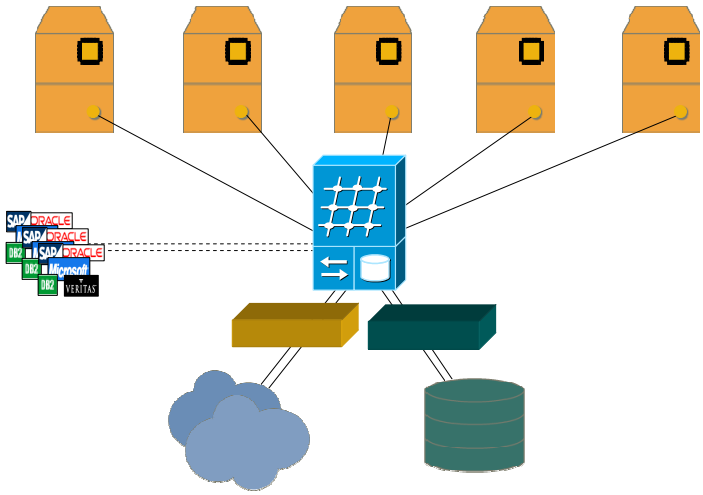
Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Evolution of the Data Center
Network and Storage Virtualization



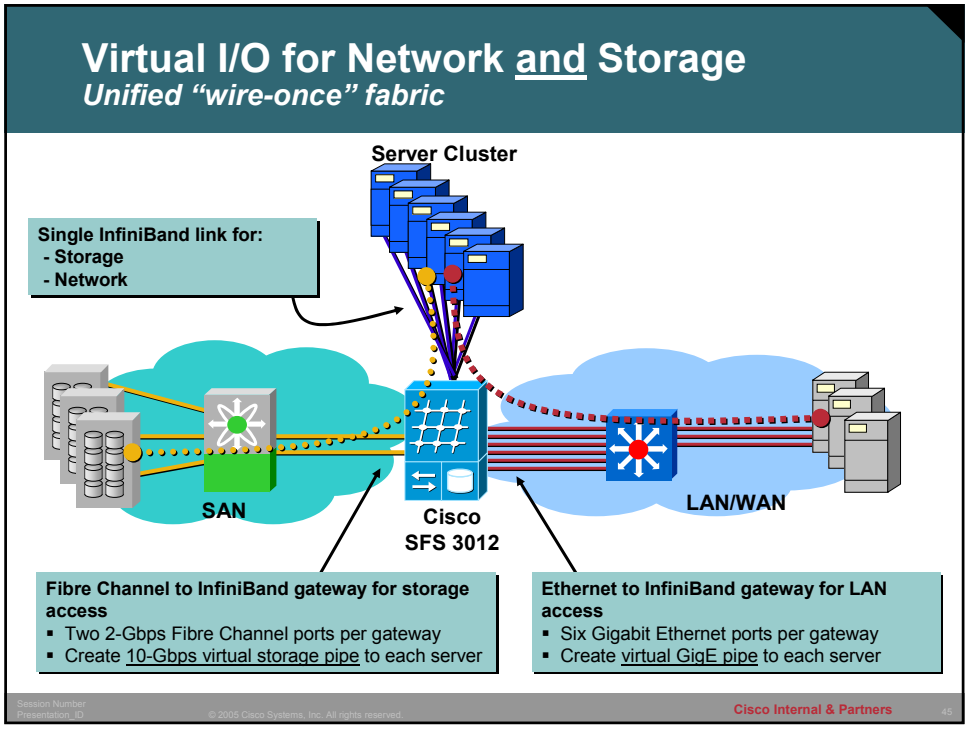
© 2005 Cisco Systems, Inc. All rights reserved.

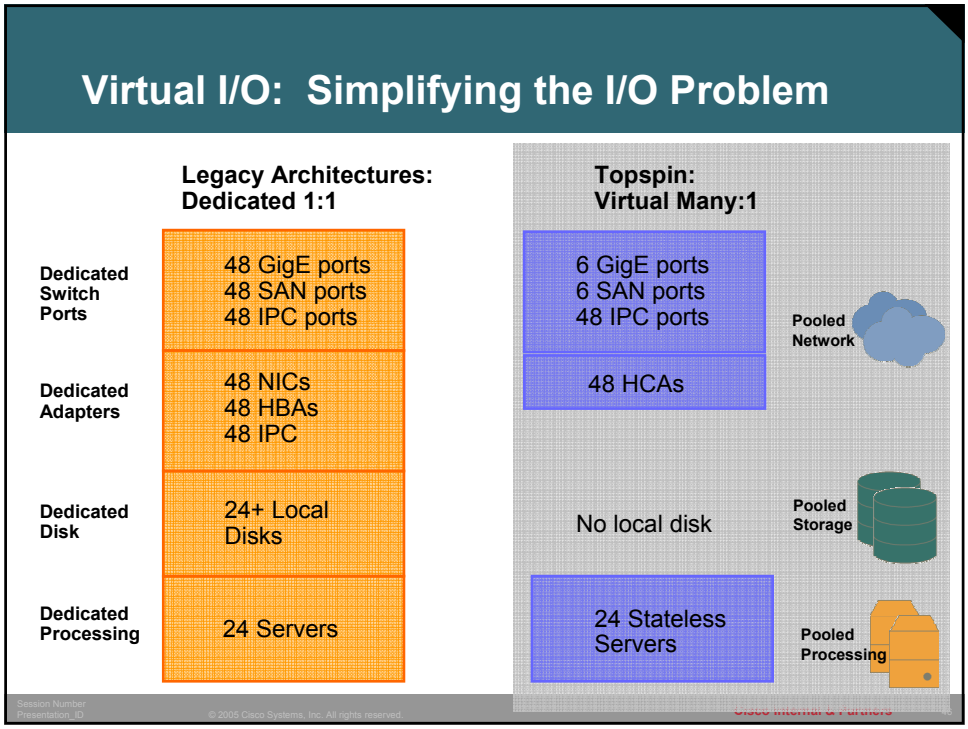
Evolution of the Data Center
Server Virtualization - The Server Switch



© 2005 Cisco Systems, Inc. All rights reserved.

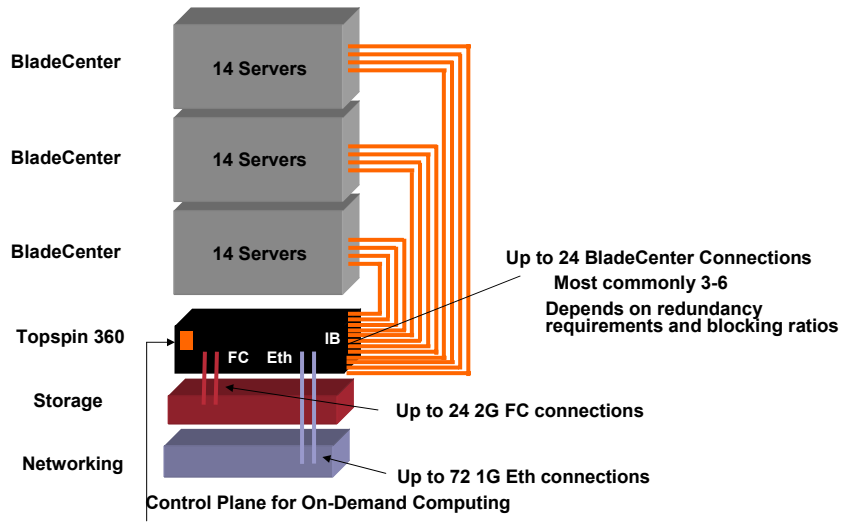
Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr


Virtual I/O for Network and Storage
Unified "wire-once" fabric
Server Cluster
Single InfiniBand link for:
- Storage
- Network
SAN
Cisco
SFS 3012
LAN/WAN
Fibre Channel to InfiniBand gateway for storage access
Two 2-Gbps Fibre Channel ports per gateway
Create 10-Gbps virtual storage pipe to each server
Ethernet to InfiniBand gateway for LAN access
Six Gigabit Ethernet ports per gateway
Create virtual GigE pipe to each server
Cisco Internal & Partners



Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

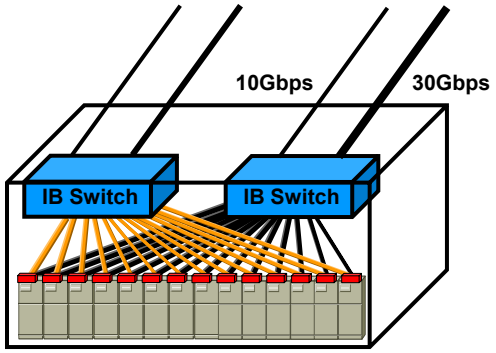
Scaling Massive I/O for the BladeCenter



Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners
47

Integrated InfiniBand for Blade Servers

- Integrated 10Gbps InfiniBand switches provide unified “wire-once” fabric
- Optimize density, cooling, space, and cable management.
- Virtual I/O provides shared Ethernet and Fibre Channel ports across blades and racks
- Option of integrated InfiniBand switch (ex: IBM BC) or pass-thru module (ex: Dell 1855)



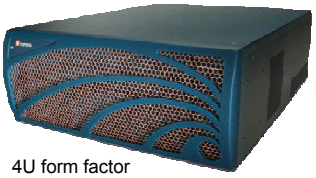
Blade Chassis with InfiniBand Switches

HCA

Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners
48

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

SFS 3012 Multifabric Server Switch



4U form factor

2 InfiniBand Switch Modules
24 10Gbps InfiniBand ports
12 hot plug expansion slots
- 2-port 2Gbps Fibre Channel
- 4-port Gigabit Ethernet

- Embedded subnet manager
- Redundant power and cooling
- Redundant control

- Dual 4X IB to each slot, dual 12 4x ports
- Console and modem ports
- Ethernet-based management port

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand-to-Ethernet Gateway Overview

- Act like L2 bridge between IB and Ethernet
- Bridge group is the main forwarding entity
- Bridge group has two bridge ports Ethernet and IPoIB
  - one VLAN to one IB partition
- Ethernet bridge port can be tagged or untagged
- Ethernet bridge port can aggregate up to 6 ports

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand-to-Ethernet Gateway Features

- IP-Only protocols
- 802.1Q VLAN support
- Link aggregation
- IPv4 multicast support
- Loop protection
- Ethernet jumbo frames up to 9k (Ingress only)
- IP fragmentation
- High Availability

© 2005 Cisco Systems, Inc. All rights reserved.

InfiniBand-to-Ethernet Gateway Performance

- Six 1 Gb/s Ports
- FPGA Based Packet Forwarding Engine
- 11 Million PPS Aggregate
- Hardware Multicast Mapping

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Ethernet Gateway - Multicast Support

- InfiniBand switches support true IB Multicast in Hardware
- InfiniBand-to-Ethernet gateways support multicast mapping in hardware.
- IB Switches use two types of Forwarding Tables:
  - Linear Forwarding Table (1 to 1 - Message In/Out)
  - Multicast Forwarding Table (1 to Many - Message In/Out)
- IB Partitions can be used to Segregate Traffic Domains
- Hardware Multicast Support means:
  - No Host Overhead for sending Multicast Messages
  - No Appreciable Latency between 1st Message & Last Message
  - No Superfluous Network Traffic
  - Multiple IB Switches in a Fabric Effectively Creates a Parallelized Multicast Delivery Mechanism (Scales Very Large, Very Fast)

© 2005 Cisco Systems, Inc. All rights reserved.

Ethernet Gateway - High Availability

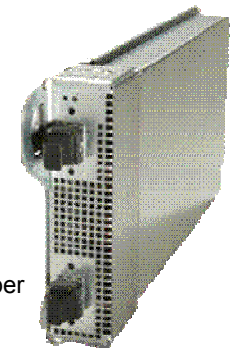
- Bridge group based redundancy
- Bridge group member of a redundancy group
- One redundancy group cover one VLAN
- Active / Passive and Active / Active modes
- Automatic fail-over and fail-back
- Uses gratuitous ARP to redirect traffic
- Redundancy group can span multiple chassis
- Proprietary redundancy protocols for address distribution and bridge group election

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

InfiniBand-to-Fibre Channel Gateway

- Ensures seamless integration with important SAN tools.
  - – Fabric-based Zoning
  - – LUN-based access controls
  - – Storage and host-based HA (Multipath) and load balancing
- Creates SAN network addresses on InfiniBand.
  - – SAN Management Tools must “see” each node.
  - – Creates “talk-through” mode with virtual WWNNs per server.
- Enables SAN Interoperability with InfiniBand.
  - – Appears as Fibre Attached Loop
  - – Proven interoperability with Brocade, McData, Qlogic, EMC, IBM, Hitachi, and more.



© 2005 Cisco Systems, Inc. All rights reserved.

Fibre Channel Gateway Performance

- Current SW Gateway:
  - – Two 2 Gb/s FC Ports
  - – Appears as Fibre Attached Loop Device
  - – 40,000 IOPs
- Next-Gen FPGA “Newton” Gateway
  - – Four 4 Gb/s FC Ports
  - – Appears as E-Port
  - – 450,000 IOPs
  - – Future support for VSANs

© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Topology Transparency: How it Works

- Storage Gateway presents either:
  - Fabric Attached Loops
  - E-Port
- SCSI RDMA (SRP) Driver installs on host as normal SCSI driver. Defined by ANSI T10 standards.
- Each IB/SRP Initiator is assigned:
  - 1 FC WWNN and
  - Multiple WWPNs
- Unique WWNs allow normal zoning to work as usual.
- Storage-based load balancing works as usual.
- Enhanced multipathing and I/O consolidation

© 2005 Cisco Systems, Inc. All rights reserved.

Topspin and Top Tier Server Vendors

IBM — "IBM and Topspin Communications Forge Key Agreement"

Sun microsystems — "Sun and Topspin Partner to Deliver Grid Computing Solutions"

DELL www.dell.com — "Dell Adds Topspin Switches to High-Performance Computing Clusters"

NEC — "Topspin Selected As NEC's Strategic InfiniBand Technology Provider"

hp — "HP To Leverage Topspin Technology"

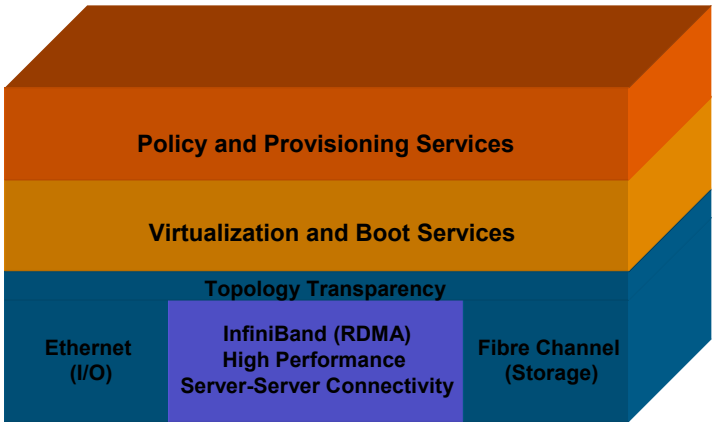
© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Server Virtualization

© 2005 Cisco Systems, Inc. All rights reserved.

VFrame Server Virtualization Framework
Building Blocks



© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

VFrame™

- **Software suite that makes the Server Switch** *programmable*
- **Three main components**

VFrame™ Embedded System Logic

Policy ingestion, interpretation, and enforcement at the server switch

VFrame™ APIs (and SDK)

Allows 3rd party (End-user Customers, Software Partners, System Vendor OEMs) management and provision tools to program and manage the server switch fabric Software Partners

VFrame™ Director

Software package disseminates policies to server switch fabric

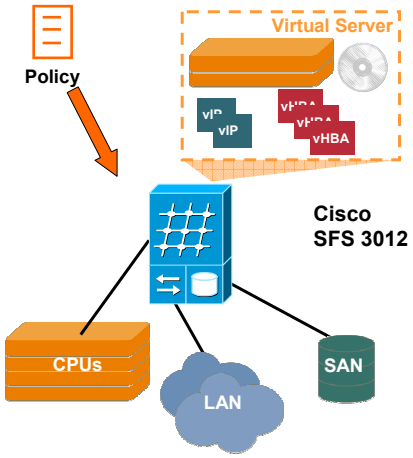
Central policy enforcement provides better system wide decision making and conflict arbitration

Can be installed on any server in the network

© 2005 Cisco Systems, Inc. All rights reserved.

Programmability
VFrame™

1) Server Switch receives policy from VFrame™ Director or 3rd party software.
2) Based on policy, Server Switch assembles the virtual server
   - Selects server(s) that meet minimum criteria (e.g. CPU, memory)
   - Boot server(s) over the network with appropriate app/os image
   - Creates virtual IPs in servers and maps to VLANs for client access.
   - Creates virtual HBAs in servers and maps to Zones, LUNs, and WWNNs for storage access



© 2005 Cisco Systems, Inc. All rights reserved.

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

How it Works
Policy Definition

A Virtual Server combines:

Everything but the physical hardware. Ex:

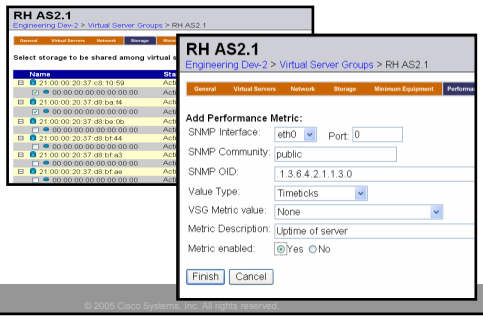
- Network Interfaces
- SAN WWNs
- Server Customization scripts

A Virtual Server Group combines:

- One or more Virtual Servers
- Shared Storage
- VLAN / SAN Zoning
- Performance Monitors
- **Policies**

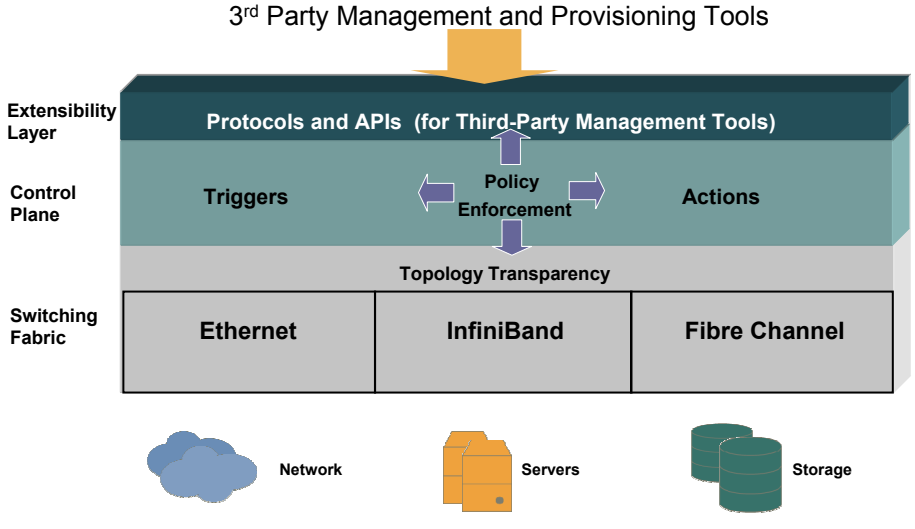
Policies Consist of:

- One or more **Trigger(s)**
  - Component Failure, Performance Metric, Scheduled Event, Custom Script
- One or more **Action(s)**
  - Add/Remove/Change Server or Group
  - Failover Server
  - Email Notification
  - Custom Script



Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners 63

Cisco SFS Architecture



Session Number
Presentation_ID
© 2005 Cisco Systems, Inc. All rights reserved.
Cisco Internal & Partners 64

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr

Copyright © 2005, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr